FIG. 1

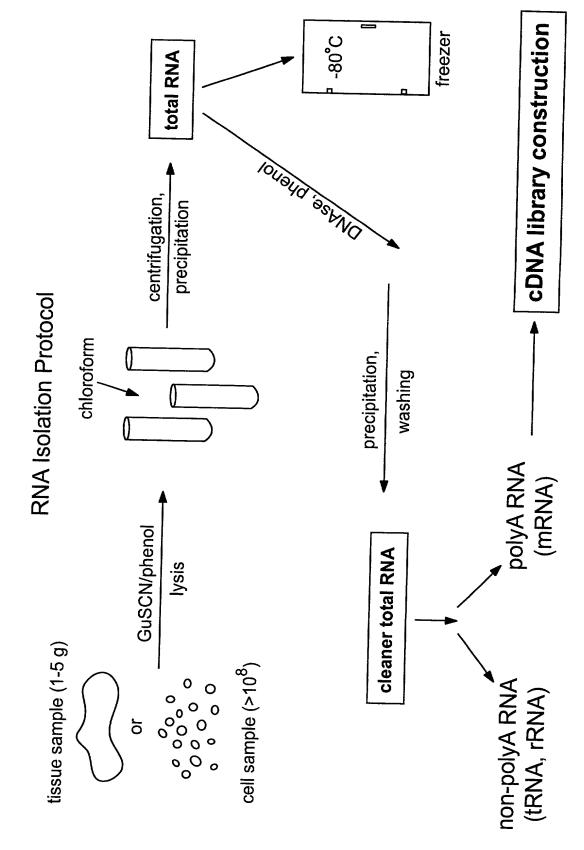
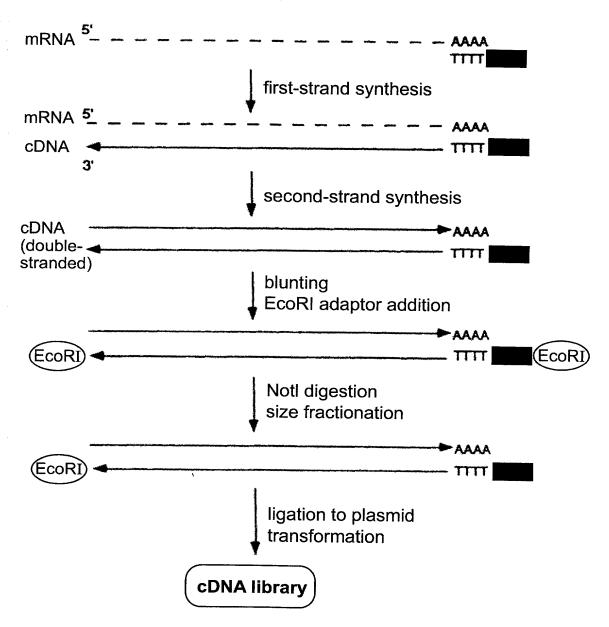


FIG. 2 cDNA Library Construction



Block 1 Sequence Editing Screens

Target Sequence Feature	Editing Method	Result
5' and 3' Vector	Dynamic Programming	Clip
PolyA Tail	Regular Expression	Clip
Sequencing Artifacts	Nearest Neighbor	Remove
Low Information	BLAST (S ≥ 90)	Mask
Contamination	BLAST (S≥90)	Remove
Repetitive Elements	BLAST (S ≥ 90)	Mask
Mitochondrial	BLAST (S≥90)	Remove
Ribosomal RNA	BLAST (S≥90)	Remove

Applicant Lincoln et al.

Title: Database for Storage and
Tarially is formula for September 19, 2001
Serial No.: To Be Assigned
Atty Docket No: 06514069CON

FIG. 4

Expected

Aberrant

	56	26	56	26	224
Н	14	4	14	14	26
ග	14	14	14	14	56
ပ	14	14	14	14	26
A	14	14	14	14	56
	A	ပ	Ø	_	
				-	
					,
	56	26	26	56	224
-	~			32 56	56 224
L 5	26	26	26		1
L 5 0	8 56	8 56	8 56	32	26

A O

U H

e.g. CCCCGGGGTTTTCCCCCAAAAGGGGG....

FIG. 5

$$E = KNe^{-\lambda S}$$
 where $S = \sum_{HSP} s(a_i, b_j)$

$$s(a,a) = 5$$
 and $s(a,b) = -4$ for $a \neq b$

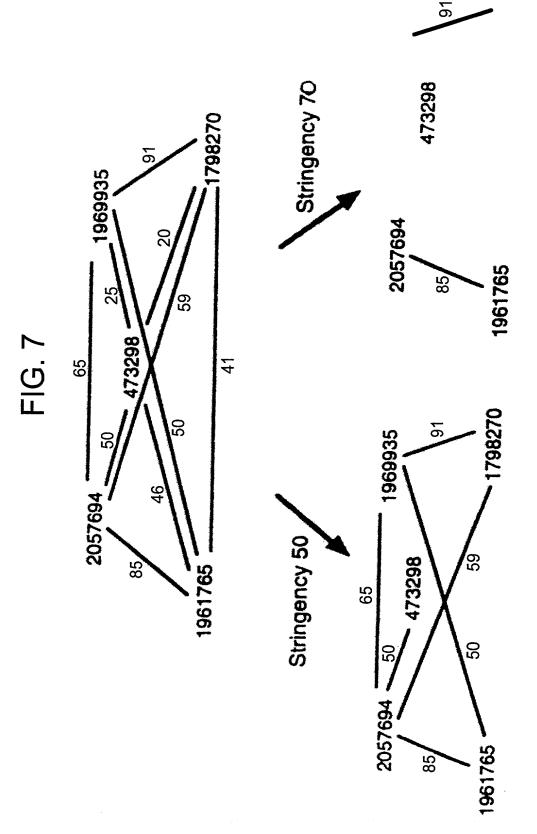
FIG. 6

Query	Match	BLAST Score	% ID	Length	Product Score
1969935 (238)	1798270	1103	99	259	91
	2057694	765	100	234	65
	1961765	600	100	245	50
	473298	318	92	261	25
473298 (261)	2057694	634	93	234	50
•	1961765	634	93	254	46
	1969935	318	92	238	25
	1798270	254	94	259	18

Alignments

1969935	
1798270	
2057694	
1961765	
473298	

Applicant Lincoln et al.
Title: Database for Storage and
Analysis circul the graph Sequences
Filing Date: September 19, 2001
Serial No.: To Be Assigned
Atty Docket No. 06514069CON



Applicant Lincoln et al.
Title: Database for Storage and
Title: Database for Storage and
Filing Date: September 19, 2001
Serial No.: To Be Assigned
Atty Docket No: 06514069CON

FIG. 8

Creating a Master Cluster

Example: Cluster - 12 (singleton), Cluster 2, and Cluster 1 all contain representative clones with PS ≥ 40 to Gene X.

Rep. Clone: 10 Cluster ID: 1 PS = 96Rep. Clone: 7 Cluster ID: 2 PS = 10075 Rep. Clone: 12 Cluster ID: -12 PS = 7575 Gene X

FIG. 9

Naming a Cluster

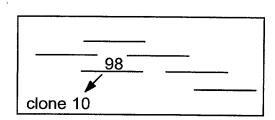
Q: Does any member of the cluster have a match to GenBank?

5 clones (solid line) match GI #1 1 clone (dashed) matches GI #2

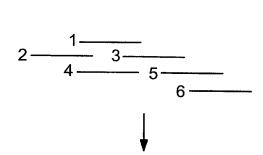
Q: Which GI is represented the most in the cluster?

A: GI #1 is represented the most.

Q: Of those clones matching GI #1, which has the highest Product Score?

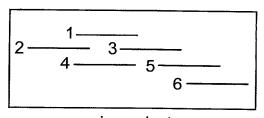


A: Clone 10 has a Product Score of 98, so the cluster is named after this representative clone.



NO

Q: Which clone has the lowest Clone ID for this cluster?



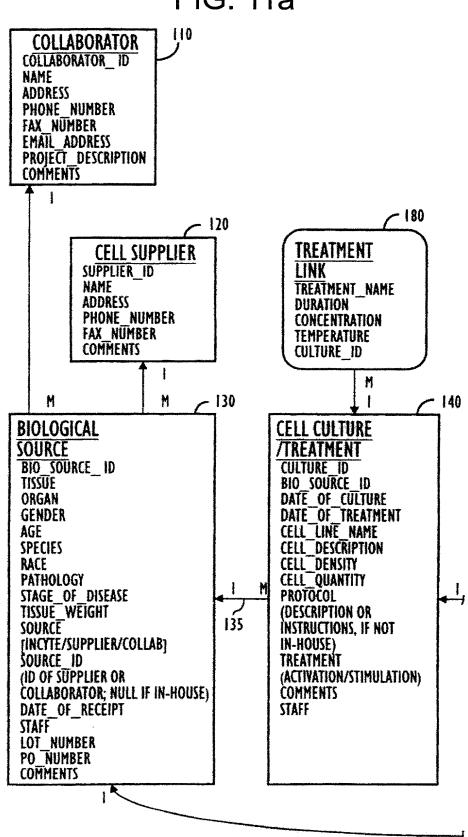
unique cluster

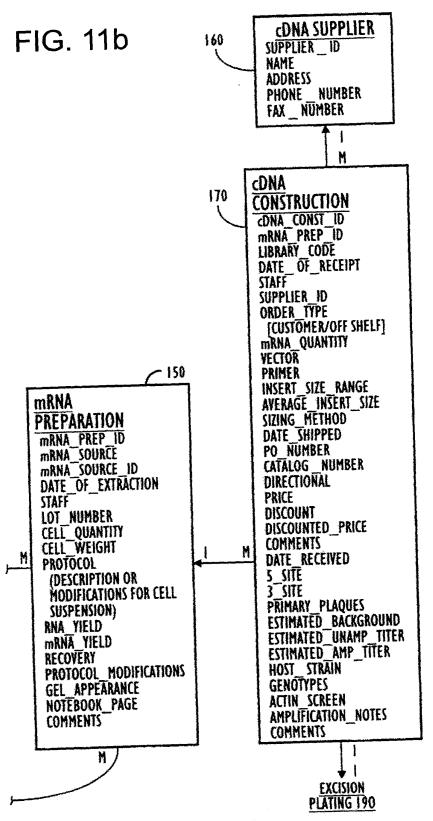
A: Clone 1 has the lowest ID, so the unique cluster is named after this representative clone.

DATABASES REFERENCE Filing Date: September 19, 2001 Serial No.: To Be Assigned Atty Docket No: 06514069CON INTEGRATED ETHERNET NETWORK DATABASE STORAGE FIG. 10 9 INTERFACE 9 COMPUTER MEMORY

Applicant Lincoln et al.
Title: Database for Storage and
Khatysis of Full Length Sequences

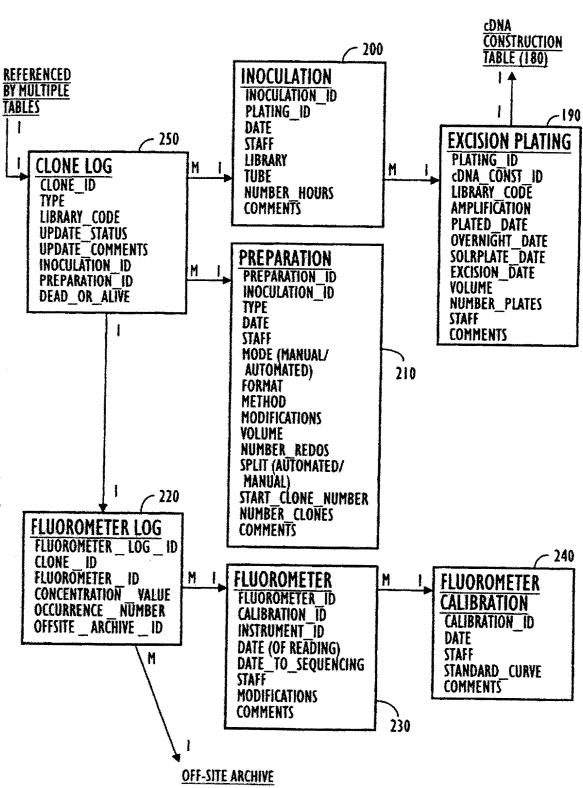
FIG. 11a

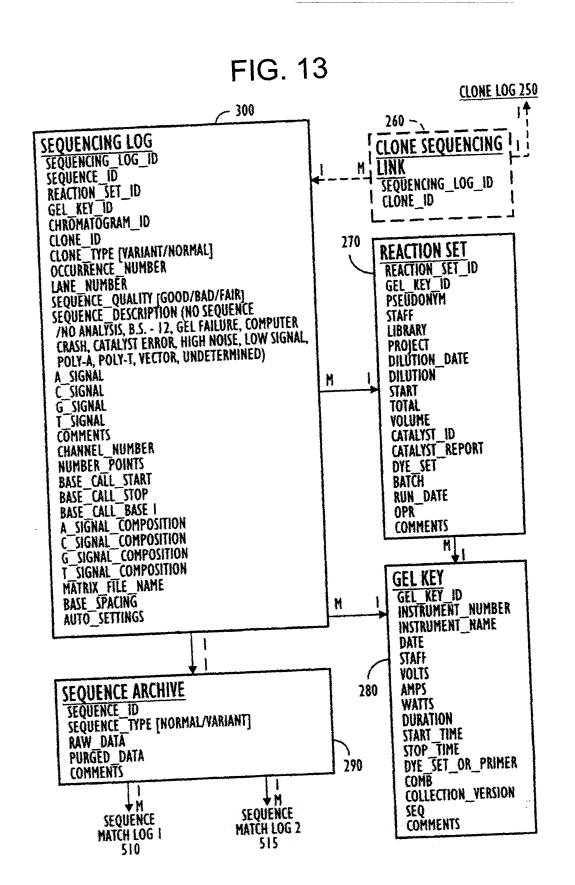




SUBSTITUTE SHEET (RULE 26)

FIG. 12





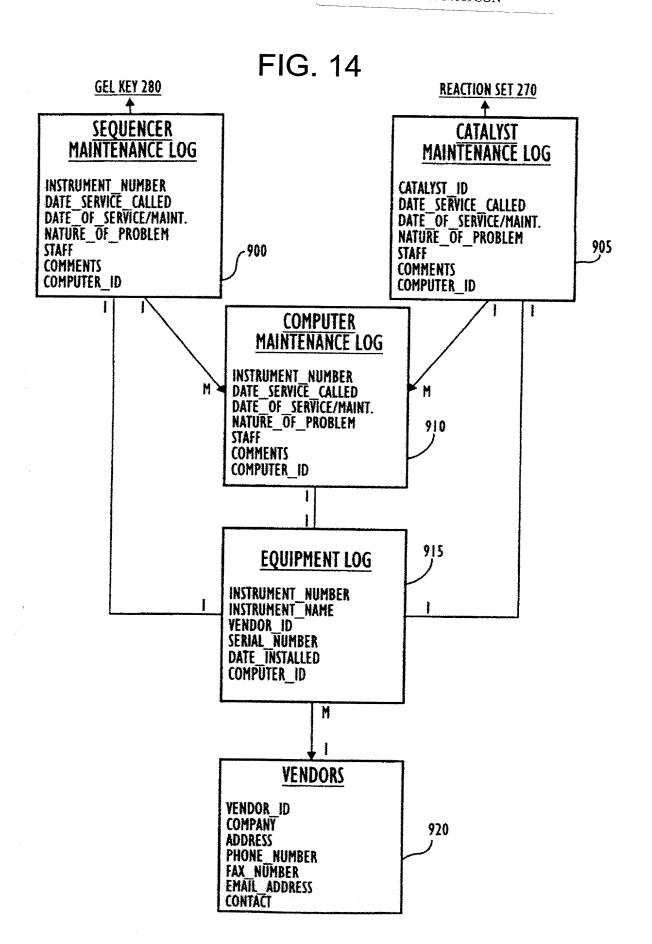


FIG. 15

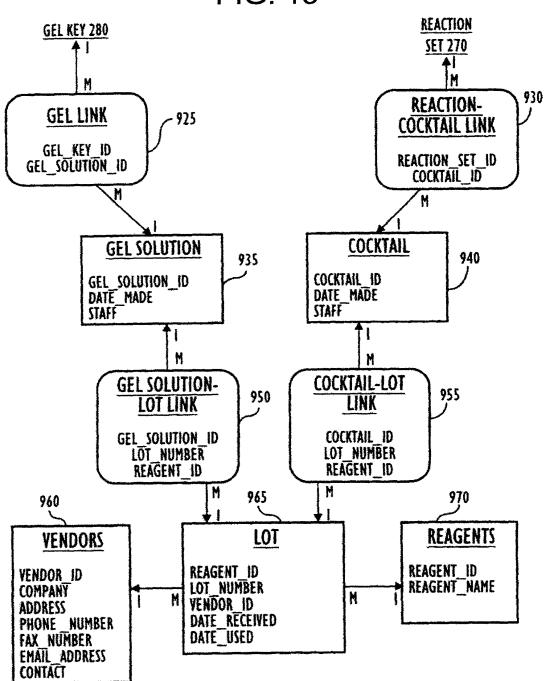
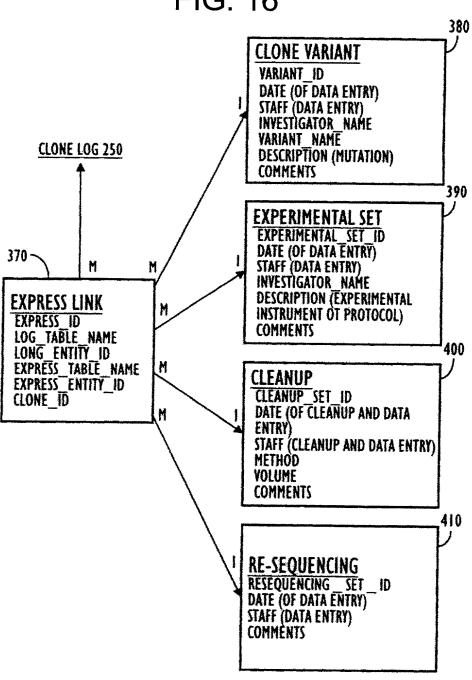


FIG. 16



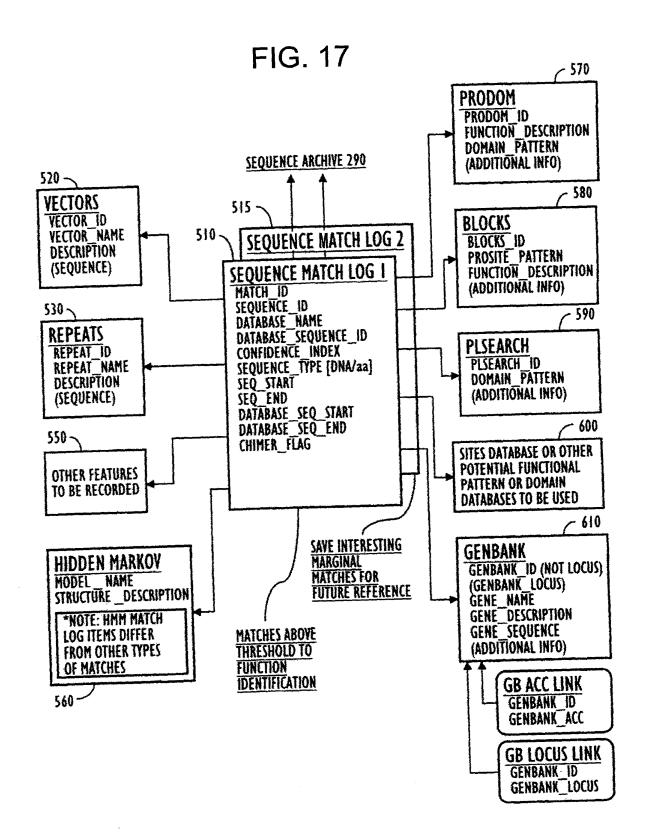


FIG. 18

